REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE

3. REPORT TYPE AND DATES COVERED
FINAL 01 JUN 91 TO 29 FEB 96

4. TITLE AND SUBTITLE
Hierarchical Learning of Complex Systems

5. FUNDING NUMBERS
AFOSR-91-0293
2304/HS 61102F

6. AUTHOR(S)
PROFESSOR DONALD A. GLASER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of California
Physics Department
Berkeley, CA 94720

8. PERFORMING ORGANIZATION REPORT NUMBER
AFOSR-TR-96-0425

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NM
110 Duncan Ave, Suite B115
Bolling AFB DC 23002-8080

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFOSR-91-0293

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE:
DISTRIBUTION UNLIMITED

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
SEE REPORT

19960822 240

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

Hierarchical Learning of Complex Systems

AFOSR 91-0293
Final Technical Report

Principal Investigator
Professor Donald A. Glaser
Project Scientist
Dr. James P. Crutchfield

Physics Department
University of California
Berkeley, California 94720 USA

Abstract

A central problem in forecasting and controlling nonlinear processes is quantifying the trade-off between available computational resources, model complexity, and prediction error. A more subtle, but important issue that strongly affects success is the choice of representation, or model class. Should one use Fourier or wavelet transforms, neural networks, hidden Markov models, or fuzzy logic, as modeling frameworks?

As a tool for answering the questions of representation dependence, brittleness, and resource requirements, we introduced hierarchical ϵ-machine reconstruction. This led to a number of detailed analyses of intrinsic computational capability in low-dimensional and spatially-extended nonlinear dynamical systems.

This Final Technical Report outlines our investigations of the computational mechanics of learning complex systems during the period beginning 1 April 1991 and ending 29 February 1996. This project was supported under AFOSR grant number 91-0293. The report reviews the activities, personnel, and research highlights and lists the published papers and those currently under review.

Contents

1 Introduction 1
2 Background 2
2.1 Computation, Dynamics, and Learning 2
2.2 ϵ-Machine Reconstruction 3
2.3 Types of Computation 4
3 Research Review 5
3.1 Thermodynamic Structure of Model Inference 6
3.2 Complexity versus Randomness 7
3.3 Semantic Information Processing: 7
3.4 Intrinsic versus Useful Computation 8
3.5 The Complexity Explosion 8
3.6 Hierarchical ϵ-Machine Reconstruction 8
3.7 Temporal Processes 9
3.8 Spatio-Temporal Processes 9
3.9 Evolving Cellular Automata to Perform Computations . . . 11
3.10 Statistical Complexity of Simple 1D Spin Systems 12
4 Invited Presentations 13
5 Research Personnel Honors 15
Bibliography 17

Introduction

This report covers the results—funded under AFOSR grant 91-0293—of our research into methods to learn models of complex nonlinear processes. The central concern in this project has been to understand how computational structure is embedded in nonlinear dynamical systems. In support of this, the initial half year — 1 April 1991 to 30 November 1991 — served to establish the personnel and computer hardware and software environment. The first full year — roughly CY1992 — saw the project gain its full momentum. Since that time we have made significant progress in most of the originally proposed problem areas, including

1. The thermodynamic structure of model inference;
2. Semantic information processing;
3. Detecting structures in nonlinear spatial systems;
4. Fluctuations in finitary stochastic processes;
5. The crucial effects of measurement distortion in temporal and spatio-temporal time series — suboptimal prediction and irreducible uncertainty;
6. Nonlinear filter design for pattern recognition and tracking;
7. A new geometric view of the infinite-state structure of recurrent hidden Markov models (HMMs);
8. An optimal entropy and statistical complexity estimation algorithm for HMMs;
9. A new measure of complexity for HMMs — the ϵ-machine dimension;
10. A new computational hierarchy for HMMs, quantified in terms of entropy rate, statistical complexity, and the ϵ-machine dimension; and
11. Evolving cellular automata to perform computations, including a discovery of how symmetry-breaking impedes the evolution to higher complexity and how embedded particles can perform decentralized computation.

During the project the AFOSR grant provided salary support for the Project Scientist and two graduate students. During its first year funds also allowed for the purchase of computing equipment.

Twenty two publications were produced during the project.[1–22] (Citations can be found in the bibliography at the Report's end.) One paper is currently under review.[20] Several reviews were published.[10,9,8] (All papers are on-line at http://www.santafe.edu/~jpc. Simply follow the "Research Communications" link.) Over thirty invited talks were presented. Three students received their Ph.D.s; two in physics and one in mathematics.[23–25] Two received fellowships. These and other personnel highlights are covered at the Report's end.

The next section provides some background needed for the review of research results, which follows immediately.

Background

Computation, Dynamics, and Learning

How does a physical system perform useful computations? At a minimum, and consistent with the underlying device physics, such a system's available physical degrees of freedom are formed into structures that constrain and guide the desired information processing. This engineering view leaves unanswered how the system's dynamical behavior supports the basic elements of computation, which include logical manipulations and the storage and transmission of information. More to the point, consider the common scientific predicament, which precedes any engineering, that the exact governing equations of motion are not available beforehand. Then, when confronted with a physical system whose behavior we wish to control or otherwise use for computation, how can the various types of embedded information processing be detected? The first step in addressing these questions is to know just what kinds of "intrinsic" computation nonlinear dynamical systems are capable of supporting. The second is to have constructive methods for inferring from a system's observable behavior the mechanisms that underlie useful computation.

This project was designed to answer these questions and also to automating *how* they can be answered. The questions themselves point to a wide range of issues, from the engineering of nanoscale processes for computation to investigations of basic scientific principles concerning the physical limits of computation, control, and modeling.

The research program we chose to answer these questions is premised on several observations. First, no globally stable, robust computation occurs in linear systems. Even the storage of a single bit of information requires nonlinearity. Second, nonlinear dynamics, both in its mathematical development over the last several decades and as seen in the exploding number of applications, provides a principled investigation of nonlinear structures underlying complex behavior.[26-29] Finally, adaptive learning, if it is anything, is a computational process, relying on information storage, decision making, and general information processing to build internal representations.

The conclusions are immediate. At the most basic level, nonlinearities underlie computation and learning however they might be instantiated in dynamical systems. We would go even further to claim that nonlinearity is essential for compact representations and efficient adaptation. Not surprisingly, the shear difficulty of problems in this area stems directly from the nonlinearities involved. The important question that remains, in any case, is how to use nonlinearity to design dynamical systems that implement a given information processing task.

Our approach to intrinsic computation and the dynamics of learning mechanisms has been intentionally interdisciplinary, combining nonlinear dynamics, statistical mechanics, information theory, and computation theory. Answers to the fundamental problems posed by intrinsic computation and by learning and generalization require this. We believe we have identified the key elements in these fields that form a basis on which to found a principled approach to analyzing intrinsic computation and to understanding how adaptive learning occurs. The combination of rigorous results and constructive implementation that we have demonstrated to date strongly suggests that the next few years will yield a corresponding range of practical applications.

The following section briefly recalls our approach to learning computational models of nonlinear processes and some related results. Several paragraphs outline two different notions of computation in dynamical systems. The project's results are then reviewed.

ϵ-Machine Reconstruction

We introduced a quantitative measure of structural complexity that reflects the intrinsic computation in nonlinear and chaotic dynamical systems.[30] For a given physical process, a computationally equivalent machine — referred to as an ϵ-machine and denoted M — can be reconstructed from a single time series. (See Table 1 for a pictorial summary of ϵ-machine reconstruction.) The technique is quite general and applies directly to the modeling task for forecasting temporal or spatio-temporal data series. The resulting minimal machine's structure indicates the inherent information processing of the original physical process. From it one can estimate information transmission, storage, and production, as well as computational properties. Our measure of structural complexity — the statistical complexity $C_\mu(M)$ — is the ϵ-machine's informational size: literally, the amount of information stored by the process. The machine states are associated with historical contexts, called morphs, that are optimal for forecasting. The simplest (topological) representation of an ϵ-machine at the lowest (nontrivial) computational level is as a stochastic deterministic finite automaton. These are displayed in the form of labeled directed graphs (l-digraphs). (Examples are shown in Table 1.) The full reconstruction, though, captures the (measure) probabilistic properties of the data stream at different computational levels. Our complexity measure unifies a number of disparate attempts to describe the information processing of nonlinear dynamical systems.[31–38]

ϵ-machine reconstruction provides a practical answer to a basic question — How does one measure the intrinsic computational properties of physical, chemical, and biological processes? This question has broad physical and engineering importance. In our framework it is answered by inferring an ϵ-machine from a data stream generated by the system under study. The resulting machine is the unique and minimal stochastic

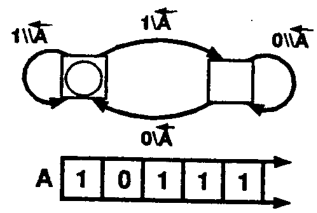
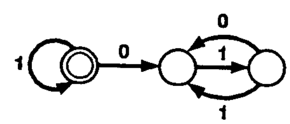
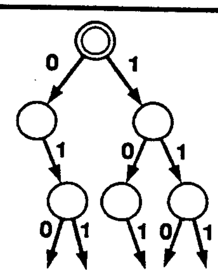
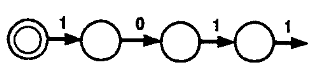
Level	Model Class	Machine	Model Size, if class is appropriate	Equivalence Relation
...	...	...	...	
3	String Production		$\mathcal{O}(\lVert\mathbf{V}\rVert + \lVert\mathbf{E}\rVert + \lVert\mathbf{P}\rVert)$	Finitary-Recursive Conditional Independence
2	Finite Automata		$\mathcal{O}(\lVert\mathbf{V}\rVert + \lVert\mathbf{E}\rVert)$	Conditional Independence
1	Tree		$\mathcal{O}\left(\lVert\mathcal{A}\rVert^{D}\right)$	Block Independence
0	Data Stream		m	Measurement

Table 1 A causal time-series modeling hierarchy. The data stream itself is the lowest level. From it a tree of depth D is constructed by grouping sequential measurements into recurring subsequences. The next level models, finite automata (FA) with states $\mathbf{V}$ and transitions $\mathbf{E}$, are reconstructed from the tree by grouping tree nodes. The last level shown, string production machines (PM), are built by grouping FA states and inferring production rules $\mathbf{P}$ that manipulate strings in register $\mathbf{A}$.

automaton consistent with the given data. In different settings, ϵ-machines can be shown to be equivalent to probabilistic versions of any one of a number of different automaton types, such as a deterministic finite automaton, a stack automaton, a queue automaton, or a cellular transducer.[39-41,4]

Types of Computation

It is essential to distinguish at least two types of information processing: "useful" computation and "intrinsic" computation.

The most common meaning of "performing a computation" is that a dynamical system carries out some "useful" information processing task. Here, the equations of motion are interpreted as the "program" and the initial state is interpreted as the "input". The system runs for some specified time until it reaches a "goal" state at which it detects the task's completion. This final condition must be relatively easy to detect. It might be, for example, a fixed point state. In any case, some portion of its configuration is interpreted as the "output". When viewed in this way, there is a correspondence between the computation and the orbit in the system's state space. Examples of dynamical systems performing useful computational tasks include integrating a differential equation on an analog computer to estimate π, using a cellular automaton to generate the n^{th} row of Pascal's triangle, performing image edge enhancement with either video feedback[42] or an oscillating chemical reaction in a petri dish,[43] running a recurrent neural network to "recover" a picture from an initially corrupted version, and — considering the notion of a dynamical system rather broadly — using a pinhole lens to estimate the Fourier transform of an image. (Discussion of useful discrete computation in cellular automata can be found in Ref. [13].)

A second meaning of computation in a dynamical system involves interpreting its behavior or, more properly, the orbits it can generate, as a type of "intrinsic" computation. Here computation is not the transformation of an input to produce a "useful" output. Rather, it is measured in terms of elementary information processing structures — memory, information production, information transfer, logical operations, and so on. In other words, intrinsic computation in a dynamical system is an intrinsic property of its behavior that can be measured by an observer just as (say) the dimension or entropy rate of the system's attractor can be estimated.[44] Intrinsic computation can be detected and quantified without reference to any specific "useful" computation performed by the dynamical system in question. In measuring it one looks at the typical information processing over the whole state space or large subsets of it. The equations of motion in this view are thought of as being the computational device in the sense that they determine the constraints that guide the flow of information in the state space. This notion of intrinsic computation was developed in Refs. [30,45,7].

Research Review

The project's central interest has been detecting computation in nonlinear dynamical systems. The approach emphasizes inferring models of nonlinear dynamical systems as a method to elucidate a system's intrinsic computational capabilities: the mechanisms by which and the rates at which the system produces information. We developed the statistical complexity C_μ to measure the former in terms of memory capacity; Shannon entropy rate h_μ is used to measure the latter — it is an indicator of randomness. During the project we've focused both on the theory of intrinsic computation and its practical application. On the theoretical side we've demonstrated that (i) in a number of cases a dynamical system's intrinsic computation is an upper bound on its usable computation, (ii) measurement distortion can result in simple processes appearing infinitely complex, and (iii) there are general principles — codified in our hierarchical reconstruction algorithms — for inferring more powerful model classes from lower level representations. The latter is the key step in detecting qualitatively different types of intrinsic computation.

This report reviews the full set of subprojects supported by AFOSR. The problem areas covered are

1. Thermodynamic structure of model inference
2. Complexity versus randomness
3. Semantic information processing
4. Intrinsic versus useful computation
5. The complexity explosion
6. Hierarchical ϵ-machine reconstruction
7. Intrinsic computation in temporal processes
   a. Deterministic dynamical systems
   b. Hidden Markov models
8. Intrinsic computation in spatio-temporal processes
   a. Cellular automata
   b. Cellular transducers
   c. Evolving cellular automata to perform computations
   d. Statistical complexity of simple 1D spin systems

The following sections cover these in turn. In each, the papers resulting from AFOSR support are cited in the subsection titles.

Thermodynamic Structure of Model Inference[2]

The original ϵ-machine reconstruction method embodies the basic elements of a general method of extracting computational structure from data series.[30] It was been generalized to automatically reconstruct a hierarchy of successively more computationally capable models.[46] As a case study of this, we analyzed the computational structure embedded in a nonlinear system at the onset of chaos.[45] The techniques developed for this allow one to analyze the structure of an important class of infinite memory processes, i.e. those with infinite correlation length, in terms of context-free grammars.

The new published work shored up these results by focusing on the relationship between equilibrium thermodynamics and the inference of models for stationary processes.[2]

In this view stationary processes correspond to equilibrium phases — gas, liquid, or ice (say). The difficulty of estimating good models, which can be measured by the statistical complexity,[30] is maximized at the boundaries between these phases. The connection between phase transitions and model complexity is now on a firm theoretical foundation. Moreover, our current modeling algorithms allow for the quantitative investigation of how modeling problems increase in difficulty when the underlying data sources are close to the phase transitions.

Perhaps of more importance, via the thermodynamic approach we have found that there are significant structures in nonlinear stationary processes that are missed by equilibrium statistical mechanics. One corollary is that if a model class "over-stochasticizes" an inference problem by (say) allowing too many statistical parameters a crippling degeneracy is added to the model optimization task. The result is that much more data and compute time are required than is necessary. This over-stochasticization is typical of hidden Markov modeling (described below). In contrast, ϵ-machine reconstruction avoids such problems by finding the minimal, statistically consistent model.

Complexity versus Randomness[30,45,9,23]

We developed universal lower bounds on how a process's complexity C_μ trades-off against the rate h_μ at which it produces information. This was developed using mathematical methods from the theory of phase transitions in physics. The result gives an interesting interpretation of physical computing devices as being in a "critical" state between thermodynamic phases of (relative) order and chaos. Critical states are characterized by their long-range spatio-temporal correlations and non-Gaussian and, typically, nonstationary statistics.

Semantic Information Processing:[1]

Once one has a learning algorithm for estimating models, the question presents itself of what to do with the models. We studied this from the viewpoint of the semantics of the measurement process. The idea here is that an observer implements in some fashion the learning algorithm so that at each moment it has an internal model of the process. Then we asked what does a given measurement mean? A quantitative answer was framed in terms of the statistical complexity of the internal model, rather than the entropy rate of the source. The result is a clear statement of the semantic content of measurements in way that does not involve subjective elements. The key step in this was the definition of a quantitative measure of meaning of individual measurements.

The longer term goal to which this contributes is understanding how nonlinear interacting systems spontaneously evolve semantics.

Intrinsic versus Useful Computation[9]

We now have a number of examples that illustrate that the intrinsic computational capacity C_μ of a dynamical system is an upper bound on the useful computation the dynamical system can perform. The examples come from both low-dimensional continuous-state and spatially-extended dynamical systems.

The Complexity Explosion[4,8,6]

We have analyzed a number of situations in which an injudicious choice of measurement discretization can make a process appear vastly more complex than it is. The examples come from hidden Markov models, discrete spatial automata, and from continuum-state dynamical systems. We have related the complexity explosion to indeterminism — the latter being used in the automata-theoretic sense — induced by the measurement process. The complexity explosion is a general phenomenon that has significant implications for developing optimal prediction and learning algorithms. With limited computational resources, for example, a process will appear more random than it is. Hierarchical machine reconstruction is our proposed solution to this problem.

Hierarchical ϵ-Machine Reconstruction[6,46]

How can an "inappropriate" choice of representation be improved in the statistical inference of models? (Note that this is a rephrasing of the problem of detecting computation in nonlinear processes.) As an answer to this very general and common scientific problem we have extended our original ϵ-machine reconstruction method to learn successively more powerful representations. The method looks at a series of increasingly-accurate models for a process. If the model size — the statistical complexity C_μ — grows without bound a new notion of "causal" machine state is inferred that captures the regularity in the *change* from one model to the next in the increasing-accuracy series. This step is referred to as "innovation". It has been analyzed in some detail for the transition (i) from finite state machines to nested stack automata, (ii) from stochastic finite state machines to infinite state stochastic machines, and (iii) from cellular automata to cellular transducers. As such hierarchical reconstruction has played a key role in our development of new computational hierarchies for probabilistic automata and for spatio-temporal processes. It also provided the key step to our demonstration of the increased computational capability at phase transitions.

Intrinsic versus Useful Computation[9]

We now have a number of examples that illustrate that the intrinsic computational capacity C_μ of a dynamical system is an upper bound on the useful computation the dynamical system can perform. The examples come from both low-dimensional continuous-state and spatially-extended dynamical systems.

The Complexity Explosion[4,8,6]

We have analyzed a number of situations in which an injudicious choice of measurement discretization can make a process appear vastly more complex than it is. The examples come from hidden Markov models, discrete spatial automata, and from continuum-state dynamical systems. We have related the complexity explosion to indeterminism — the latter being used in the automata-theoretic sense — induced by the measurement process. The complexity explosion is a general phenomenon that has significant implications for developing optimal prediction and learning algorithms. With limited computational resources, for example, a process will appear more random than it is. Hierarchical machine reconstruction is our proposed solution to this problem.

Hierarchical ϵ-Machine Reconstruction[6,46]

How can an "inappropriate" choice of representation be improved in the statistical inference of models? (Note that this is a rephrasing of the problem of detecting computation in nonlinear processes.) As an answer to this very general and common scientific problem we have extended our original ϵ-machine reconstruction method to learn successively more powerful representations. The method looks at a series of increasingly-accurate models for a process. If the model size — the statistical complexity C_μ — grows without bound a new notion of "causal" machine state is inferred that captures the regularity in the *change* from one model to the next in the increasing-accuracy series. This step is referred to as "innovation". It has been analyzed in some detail for the transition (i) from finite state machines to nested stack automata, (ii) from stochastic finite state machines to infinite state stochastic machines, and (iii) from cellular automata to cellular transducers. As such hierarchical reconstruction has played a key role in our development of new computational hierarchies for probabilistic automata and for spatio-temporal processes. It also provided the key step to our demonstration of the increased computational capability at phase transitions.

Temporal Processes

There are two broad process classes — temporal and spatio-temporal — that we can analyze from the viewpoint of intrinsic computation. This section and the following one list some results.

Deterministic dynamical systems[45,8,6,23]

The processes defined by these models give us an arena — continuous-state systems — in which to test the applicability of our computation theoretic approach in the sense that they are reasonable models for a wide range of natural processes. We are currently working on improving the estimation of fluctuation properties[11] for these processes via generating and nongenerating partitions of their state spaces.

Hidden Markov models[8,6,25]

The class of processes of interest here has been variously labeled by the disciplines that have studied them as hidden Markov models (HMMs), functions of a Markov chain, stochastic nondeterministic automata, and communication channels. The main results to date are as follows.

1. We have an on-line algorithm for optimally predicting HMM time series. This gives an efficient way for estimating the entropy rate h_μ and the statistical complexity C_μ for HMMs.
2. We have a new computational classification for HMMs: Denumerable Stochastic Automata, Fractal Stochastic Automata, and Continuum Stochastic Automata. The inference resources for each type increases dramatically (and in the order just specified).
3. We have a new measure of the difficulty of predicting HMM time series — the ϵ-machine dimension $d_{\epsilon M}$. We now are studying the relationship between this quantity, the entropy rate, and the statistical complexity for HMMs.

Spatio-Temporal Processes

Cellular automata[3,7,47,5,24,17]

The ϵ-machine reconstruction procedure[30] was adapted to spatially-extended systems in order to quantify the complexity of patterns.[47] The result is the reconstruction of space-time machines that describe the complexity of pattern evolution over ensembles of space-time paths. The existing statistical mechanical and thermodynamic description of ϵ-machines carries over directly to give quantitative measures of entropy and complexity densities. The corresponding thermodynamics allows one to investigate mixed-phase systems. The phases, for example, are associated with time-invariant, spatially-homogeneous "domains" and describe different computational capacities.

One practical application of reconstructed space-time machines is to use them to "nonlinearly filter" time-dependent patterns to detect propagating coherent structures. We have made quite of bit of progress by applying these techniques to cellular automata (CA) — spatially-extended systems that are discrete in space, in time, and in local state value.

Using our methods unpredictable patterns generated by CA can be decomposed with respect to a turbulent, positive entropy rate pattern basis. The resulting patterns uncover significant structural organization in a CA's dynamics and information processing capabilities. In [7] we illustrated the decomposition technique by analyzing a binary, range-2 cellular automaton having two invariant chaotic domains of different complexities and entropies. Once they were identified, the domains were seen to organize the CA's state space and to dominate its evolution. Starting from the domains' structures, we showed how to construct a finite-state transducer that performs nonlinear spatial filtering such that the resulting space-time patterns reveal the domains and the intervening walls and dislocations. To show the statistical consequences of domain detection, we compared the entropy and complexity densities of each domain with the globally averaged (nonstationary) quantities. A more graphical comparison was also used: difference patterns and difference plumes which trace the space-time influence of a single-site perturbation. We also investigated the diversity of walls and particles emanating from the interface between two adjacent domains.

We have now firmly established the general applicability of the formal development of qualitative dynamics of spatial systems introduced in [47]. Indeed, we expect much further theoretical progress and a range of technological applications to follow.

Cellular transducers[4]

Cellular automata (CA) form one of the simplest model classes for spatial pattern generating processes. CA have been proposed as models of pattern formation in natural systems.[48-51] Verifying this has been largely a matter of comparing CA behavior, as revealed in (say) space-time diagrams, snapshots of spatial patterns, and various macroscopic statistics produced during computer simulation, with natural patterns. More recently, several authors suggested that effective CA equations of motion, consisting of a look up table that maps neighborhood templates to next site value, could be inferred from pattern data time series.[52-56] The learning paradigm employed, however, did not take into account the effect of measurement distortion common in obtaining experimental data. We showed in [4] that the latter can have a fundamental effect on the success of CA estimation, in particular, and spatial modeling, generally.

The foremost cause of this is that measurements are only indirect representations of a process's internal states. One practical consequence of this basic physical fact is that cellular transducers (CT) — a new class of models introduced in [4] which explicitly account for the measurement process — rather than cellular automata (CA), should be used as the computational model class for reconstructing the spatio-temporal dynamic from pattern data series. The latter includes data generated by discrete-state systems and the spatio-temporal symbolic dynamics of continuum-state extended systems, such as map lattices,[57] oscillator chains, and partial differential equations. The main difficulty is that estimated CA look up tables (LUTs) misrepresent the dynamics even if the observed behavior was generated by a deterministic process with finite local memory. Examples of nearest-neighbor binary-alphabet CT with two local states were given in [4] that require an infinite CA LUT for their deterministic dynamics to be (i) effectively reconstructed, (ii) approximately reconstructed, and (iii) not reconstructed at all. In these cases any estimated CA is stochastic and, as such, fails to capture obvious spatio-temporal structure. This leads to an overestimation of the degree of intrinsic randomness underlying the spatial data series.

One of the larger implications for research directions in nonlinear modeling is that a concerted effort is required in understanding the computational structure of nonlinear dynamical systems. Without a principled understanding of the effects of incorrect model class assumptions, scientists and engineers will fail in their ability to extract significant structures from data. The consequences for suboptimal pattern recognition are clear.

Evolving Cellular Automata to Perform Computations[12–16,19,18,21,22]

The study of how nonlinear dynamical systems support computation involves a number of issues and concepts from different disciplines. In particular, How does computational capability relate to dynamical behavior? How predictive of computational capability are statistical and information theoretic characterizations of behavior?

There has been a good deal of interest in the basic questions surrounding these questions under the rubric of "computation at the edge of chaos". The idea being that dynamical systems at the onset of chaos are the most computationally capable. The first concrete results were presented by us in [30,45]. Since that time, however, there has been a great deal of speculation and resulting confusion about our basic results stemming from studies of cellular and Boolean automata. This has been exacerbated by the further hypothesis that evolutionary systems will evolve spontaneously to the "edge of chaos" and so become more complex.

We engaged in a substantial effort to clarify the basic issues revolving around the questions of evolution, behavior, and computation. In [12,14], we presented results from an experiment similar to one performed by Packard,[58] in which a genetic algorithm (GA) is used to evolve cellular automata (CA) to perform a particular computational task. Packard's original study examined the frequency of evolved CA rules as a function of Langton's λ parameter,[59] and interpreted the results of his experiment as giving evidence for the following two hypotheses:

1. CA rules able to perform complex computations are most likely to be found near "critical" λ values, which have been claimed to correlate with a phase transition between ordered and chaotic behavioral regimes for CA; and
2. When CA rules are evolved to perform a complex computation, evolution will tend to select rules with λ values close to the critical values.

Our extensive experiments produced very different results. We concluded that the interpretation of the original results is not correct. In [13] we also reviewed and clarified issues related to λ, dynamical-behavior classes, and computation in CA.

The main constructive results of our study was identifying the emergence and competition of computational strategies and analyzing the central role of symmetries in an evolutionary system. In particular, we demonstrated how symmetry breaking can impede the evolution toward higher computational capability. We also demonstrated how a genetic algorithm can discover CA that use embedded particles and their interactions to perform decentralized computation.

We feel that with these constructive results and the critique of the earlier work we are now in a very strong position to analyze the detailed mechanisms that govern the genetic algorithm's success — or lack thereof. The results we anticipate from this development should have major implications for automatically programming distributed computational systems, like CA.

Statistical Complexity of Simple 1D Spin Systems[20]

Given the central role developing for statistical physics methods in the analysis of complex and learning systems, it is extremely important to understand how our computational quantifiers relate to existing observables in statistical mechanics and thermodynamics. We derived exact results for two complementary measures of spatial structure generated by 1D spin systems with finite-range interactions. The first, excess entropy, measures the apparent spatial memory stored in configurations. The second, statistical complexity, measures the amount of memory needed to optimally predict the chain of spin values in configurations. It turns out that these statistics capture distinct properties and are different from existing thermodynamic quantities. This suggests new ways to detect computational structure in thermodynamic (large-scale) systems.

Invited Presentations

The following is a list of invited talks given by members of our group during the report period. Contributed presentations and posters are not cited. Below JPC refers to Jim Crutchfield and JEH to Jim Hanson.

1. JPC, *Computation in Chaos*, Lecture course, sponsored by the Fuzzy Logic Systems Institute, 20 - 21 September 1991, Fukuoka, Japan, and 24 - 25 September 1991, Tokyo, Japan; Colloquium, Center for Complex Systems Research, Beckman Institute, University of Illinois, Champaign-Urbana, 11 October 1991; Institute for Scientific Computing Research, Lawrence Livermore National Laboratory, Livermore, California, 2 July 1992; Colloquium, Interval Research, Inc., Palo Alto, California, 8 June 1993.
2. *The Attractor-Basin Portrait of a Cellular Automaton*, Colloquium, Santa Fe Institute, Santa Fe, New Mexico, 1 March 1991.
3. *Discovering Coherent Structures in Nonlinear Spatial Systems*, at the Applied Physics Laboratory Symposium on **Nonlinear Dynamics of Ocean Waves**, Johns Hopkins University, Maryland, 30 — 31 May 1991.
4. *Computational Mechanics: toward a physics of complexity*, lecture course, Beckman Institute, University of Illinois, Urbana-Champaign, November - December 1991.
5. *Dynamics and Model Inference* and *The Semantics of Mechanical Systems*, Conference on **Dynamic Representations in Cognition**, Indiana University, Bloomington, Indiana, 15 and 16 November 1991.
6. JPC, *Computation in Chaos: toward a physics of complexity*, Dynamics Days, Texas, Austin, Texas, 8 - 11 January 1992.
7. JPC, *The Calculi of Emergence: Complexity as the Interplay of Order and Chaos*, Santa Fe Institute Integrative Themes Workshop, Santa Fe, New Mexico, 8 - 15 July 1992.
8. JPC, *Thermodynamics of Inference*, NATO Advanced Studies Institute **From Statistical Physics to Statistical Inference and Back**, Cargese, France, 31 August - 12 September 1992.
9. JPC, *Innovation, Induction, and Complexity*, Santa Fe Institute Workshop on **Computation, dynamical systems, and learning** Santa Fe, New Mexico, 16 - 20 November 1992.

10. JEH, *Chaotic Pattern Bases for Cellular Automata*, Santa Fe Institute Workshop on**Computation, dynamical systems, and learning** Santa Fe, New Mexico, 16 - 20 November 1992.
11. JPC, *Critical Computation and Hierarchical Learning*, Colloquium, Institute for Theoretical Physics and Synergetics, University of Stuttgart, Germany, 26 March 1993; Nonlinear Dynamics Seminar, Tokyo Institute of Technology, Tokyo, Japan, 12 April 1993.
12. JPC, *Observing Complexity and the Complexity of Observation*, Max Planck Institute sponsored workshop on **Endo-Exo Problems in Physics**, Ringberg Castle, Bavaria, Germany, 29 March - 2 April 1993.
13. JPC, *The Calculi of Emergence: Complexity as the Induction of Order and Chaos*, 36^{th} Oji International Seminar on **Complex Systems — from Complex Dynamical Systems to the Sciences of Artificial Reality**, Fujitsu Forum, Numazu City, Japan, 5 - 9 April 1993.
14. JPC, *Turbulent Pattern Bases for Spatial Systems*, APS meeting **Computational Physics 1993**, Albuquerque, New Mexico, 3 June 1993.
15. JPC, *Fluctuation Spectroscopy*, Workshop on **Fluctuations and Order: the new synthesis**, Los Alamos, New Mexico, 9 September 1993.
16. JPC, *Critical Computation, Phase Transitions, and Hierarchical Learning*, The Seventh Toyota Conference **Towards the Harnessing of Chaos**, Mikkabi, Japan, 1 November 1993.
17. JPC, *Towards a Statistical Dynamics of Genetic Algorithms*, Workshop on **Theoretical Foundations of Genetic Algorithms**, Santa Fe Institute, New Mexico, 11 - 13 January 1994.
18. JPC, *The Evolution of Emergent Computation*, **International Conference on Dynamical Systems and Chaos**, Tokyo Metropolitan University, Tokyo, Japan, 23 - 27 May 1994.
19. JPC, *Computational Mechanics: Towards a Physics of Complexity*, Two lectures presented to the Extended Workshop on **Dynamics and Complexity**, Technical University, Lisbon, Portugal, 14 and 16 September 1994; invited review presented to the Workshop on **Theory and Applications of Nonlinear Time Series Analysis**, Potsdam, 20-30 September 1995; Symposium on **Computational Issues in Learning Dynamical Systems**, AAAI Spring Meeting, Stanford University, 26 March 1996.
20. JPC, *How do Nonlinear, Time-Dependent Processes Compute?*, **Neurosciences Institute 11^{th} Summer Atelier on Theoretical Neurobiology**, La Jolla, California, 23 September 1994.
21. JPC, *Observing Complexity and the Complexity of Observation*, Joint Physics and Philosophy Colloquium, Reed College, Portland, Oregon, 9 November 1994.

22. JPC, *The Evolution of Emergent Computation*, Seminar, Mathematics Department, Reed College, Portland, Oregon, 10 November 1994; CIRES Colloquium, University of Colorado, Boulder, Colorado, 13 April 1995; International Conference on **Self-Organization of Complex Structures**, Berlin, 24-28 September 1995.
23. JPC, *How Does Nature Compute?*, Joint Colloquium, Keck Center for Integrative Neurobiology and Sloan Center for Theoretical Neurobiology, University of California, San Francisco. 15 December 1995.
24. JPC, *What is a Pattern? Discovering the Hidden Order in Chaos*, Bernard Osher Fellowship public lecture, San Francisco Exploratorium, 3 July 1996.

Research Personnel Honors

Distinguished Visiting Research Professor: The grant's project scientist Dr. Crutchfield spent Fall of 1991 as a Distinguished Visiting Research Professor at the Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign. While there he gave a series of ten lectures on Computational Mechanics, the research supported by the AFOSR grant.

Bernard Osher Foundation Fellow: Dr. Crutchfield was the principal scientific advisor for a large exhibition mounted by the San Francisco Exploratorium, a widely-known and respected science museum. The exhibition, titled **Turbulent Landscapes: The Forces That Shape Our World**, is funded by the NSF and by DOE. Its exhibits demonstrate a range of phenomena related to turbulence, pattern formation, complexity, and chaotic dynamical systems. It will tour nationally and internationally over the coming three years. It is estimated that 3 to 4 million visitors will eventually see the exhibition. Dr. Crutchfield was awarded a Bernard Osher Foundation Fellowship in 1995–1996 to continue helping the Exploratorium mount the exhibition.

Santa Fe Institute Research Professor: Dr. Crutchfield is now Research Professor at the Santa Fe Institute, where he has initiated a 5–year research program ($400K/yr) in **Computation, Dynamics, and Inference**; a research area that is a direct outgrowth of the AFOSR project. Dr. Crutchfield will be Scientific Director of the DARPA-funded ($5M) SFI program on **Foundations of Complex Adaptive Systems**.

National Research Council Post-Doctoral Fellowship: Karl Young, who was supported by the AFOSR grant for his graduate student research on computational mechanics, was awarded a two-year National Research Council Post-Doctoral Fellowship. He worked at the Space Sciences Division at NASA-Ames Research Center in Moffett Field, California, and continues to collaborate with Dr. Crutchfield on research of direct concern to the AFOSR projects. Dr. Young is currently a Research Physicist at the University of California, San Francisco, Department of Radiology, where he is developing new methods of spatial (fMRI) data analysis based on ϵ-machine reconstruction and wavelet transforms.

National Science Foundation Graduate Fellow: Dan Upper, a graduate student in the UCB Mathematics Department joined our group in Spring 1992. At the time he was supported by an NSF graduate fellowship. He is now a graduate research assistant supported by the AFOSR project. He initially worked on nonlinear time series prediction and modeling and now has focused on hidden Markov models — their computational and statistical structure and optimal algorithms for estimating their properties. His dissertation will be completed during Summer 1996.

Santa Fe Institute Post-Doctoral Fellow: Jim Hanson, who finished his Physics Ph.D. in August 1993 with AFOSR support, was selected out of a field of more than two hundred applicants for a two-year post-doctoral fellowship at the Santa Fe Institute in Santa Fe, New Mexico. Dr. Hanson has recently taken a position with IBM Thomas J. Watson Research Center, Yorktown Heights, New York, to work on communication network dynamics, control, and security.

Bibliography

1. J. P. Crutchfield. Knowledge and meaning ... chaos and complexity. In L. Lam and V. Naroditsky, editors, *Modeling Complex Phenomena*, pages 66 - 101, Berlin, 1992. Springer-Verlag.
2. J. P. Crutchfield. Semantics and thermodynamics. In M. Casdagli and S. Eubank, editors, *Nonlinear Modeling and Forecasting*, volume XII of *Santa Fe Institute Studies in the Sciences of Complexity*, pages 317 - 359, Reading, Massachusetts, 1992. Addison-Wesley.
3. J. P. Crutchfield. Discovering coherent structures in nonlinear spatial systems. In A. Brandt, S. Ramberg, and M. Shlesinger, editors, *Nonlinear Ocean Waves*, pages 190–216, Singapore, 1992. World Scientific. also appears in Complexity in Physics and Technology, R. Vilela-Mendes, editor, World Scientific, Singapore (1992).
4. J. P. Crutchfield. Unreconstructible at any radius. *Phys. Lett. A*, 171:52 - 60, 1992.
5. J. P. Crutchfield and J. E. Hanson. Attractor vicinity decay for a cellular automaton. *CHAOS*, 3(2):215–224, 1993.
6. J. P. Crutchfield. The calculi of emergence: Computation, dynamics, and induction. *Physica D*, 75:11 - 54, 1994.
7. J. P. Crutchfield and J. E. Hanson. Turbulent pattern bases for cellular automata. *Physica D*, 69:279 - 301, 1993.
8. J. P. Crutchfield. Observing complexity and the complexity of observation. In H. Atmanspacher, editor, *Inside versus Outside*, pages 235 - 272, Berlin, 1994. Springer-Verlag.
9. J. P. Crutchfield. Critical computation, phase transitions, and hierarchical learning. In M. Yamaguti, editor, *Towards the Harnessing of Chaos*, Amsterdam, 1994. Elsevier Science. Santa Fe Institute Technical Report 93-10-061.
10. J. P. Crutchfield. Is anything ever new? Considering emergence. In G. Cowan, D. Pines, and D. Melzner, editors, *Complexity: Metaphors, Models, and Reality*, volume XIX of *Santa Fe Institute Studies in the Sciences of Complexity*, pages 479 - 497, Reading, MA, 1994. Addison-Wesley. Santa Fe Institute Technical Report 94-03-011.
11. K. Young and J. P. Crutchfield. Fluctuation spectroscopy. *Chaos, Solitons, and Fractals*, 4:5 - 39, 1994.
12. M. Mitchell, P. Hraber, and J. P. Crutchfield. Revisiting the edge of chaos: Evolving cellular automata to perform computations. *Complex Systems*, 7:89 - 130, 1993.
13. M. Mitchell, J. P. Crutchfield, and P. Hraber. Dynamics, computation, and the "edge of chaos": A re-examination. In G. Cowan, D. Pines, and D. Melzner, editors,

Complexity: Metaphors, Models, and Reality, volume XIX of *Santa Fe Institute Studies in the Sciences of Complexity*, pages 497 – 513, Reading, MA, 1994. Addison-Wesley. Santa Fe Institute Technical Report 93-06-040.

14. M. Mitchell, J. P. Crutchfield, and P. T. Hraber. Evolving cellular automata to perform computations: Mechanisms and impediments. *Physica D*, 75:361 – 391, 1994.

15. J. P. Crutchfield and M. Mitchell. The evolution of emergent computation. *Proc. Natl. Acad. Sci.*, 92:10742 – 10746, 1995.

16. R. Das, M. Mitchell, and J. P. Crutchfield. A genetic algorithm discovers particle computation in cellular automata. In Y. Davidor, H.-P. Schwefel, and R. Männer, editors, *Parallel Problem Solving in Nature—PPSN III*, volume 866 of *Lecture Notes in Computer Science*, pages 344–353, Berlin, 1994. Springer-Verlag. Santa Fe Institute Technical Report 94-03-015.

17. J. E. Hanson and J. P. Crutchfield. Computational mechanics of cellular automata: An example. *Physica D*, page to appear, 1996.

18. M. Mitchell, J. P. Crutchfield, and R. Das. Evolving cellular automata to perform computations. In T. Back, D. Fogel, and Z. Michalewicz, editors, *Handbook of Evolutionary Computation*, page in press. Oxford University Press, Oxford, 1996.

19. R. Das, J. P. Crutchfield, M. Mitchell, and J. E. Hanson. Evolving globally synchronized cellular automata. In L. J. Eshelman, editor, *Proceedings of the Sixth International Conference on Genetic Algorithms*, pages 336 – 343, San Francisco, CA, 1995. Morgan Kaufmann.

20. J. P. Crutchfield and D. P. Feldman. Statistical complexity of simple 1D spin systems. *Physical Review E*, page submitted, 1996.

21. M. Mitchell, J. P. Crutchfield, and R. Das. Evolving cellular automata to perform computations: A review of recent work. In *Proceedings of the First International Conference on Evolutionary Computation and Its Applications*, page in press, Moscow, Russia, 1996. Russian Acadamy of Sciences.

22. W. Hordijk, J. P. Crutchfield, and M. Mitchell. Embedded-particle computation in evolved cellular automata. In T. Toffoli, editor, *Proceedings of the Fourth Workshop on Physics and Computation: PhysComp96*, Amsterdam, 1996. Elsevier.

23. K. Young. *The Grammar and Statistical Mechanics of Complex Physical Systems*. PhD thesis, University of California, Santa Cruz, 1991. published by University Microfilms Intl, Ann Arbor, Michigan.

24. J. E. Hanson. *Computational Mechanics of Cellular Automata*. PhD thesis, University of California, Berkeley, 1993. Published by University Microfilms Intl, Ann Arbor, Michigan.

25. D. R. Upper. *Computational Mechanics of Recurrent Hidden Markov Models*. PhD thesis, University of California, Berkeley, 1996. Published by University Microfilms Intl, Ann Arbor, Michigan.

26. J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Springer-Verlag, New York, 1983.

27. A. J. Lieberman and M. A. Lichtenberg. *Regular and Stochastic Motion*. Springer-Verlag, New York, 1983.

28. R. L. Devaney. *An Introduction to Chaotic Dynamical Systems*. Addison-Wesley, Redwood City, California, 1989.

29. S. Wiggins. *Global Bifurcations and Chaos: analytical methods*. Springer-Verlag, New York, 1988.

30. J. P. Crutchfield and K. Young. Inferring statistical complexity. *Phys. Rev. Let.*, 63:105 - 108, 1989.

31. J. P. Crutchfield and N. H. Packard. Symbolic dynamics of one-dimensional maps: Entropies, finite precision, and noise. *Intl. J. Theo. Phys.*, 21:433, 1982.

32. J. P. Crutchfield and N. H. Packard. Symbolic dynamics of noisy chaos. *Physica*, 7D:201 - 223, 1983.

33. S. Wolfram. Computation theory of cellular automata. *Comm. Math. Phys.*, 96:15, 1984.

34. R. Shaw. *The Dripping Faucet as a Model Chaotic System*. Aerial Press, Santa Cruz, California, 1984.

35. C. H. Bennett. On the nature and origin of complexity in discrete, homogeneous locally-interacting systems. *Found. Phys.*, 16:585, 1986.

36. C. P. Bachas and B. A. Huberman. Complexity and relaxation of hierarchical structures. *Phys. Rev. Let.*, 57:1965, 1986.

37. P. Grassberger. Toward a quantitative theory of self-generated complexity. *Intl. J. Theo. Phys.*, 25:907, 1986.

38. H. Pagels and S. Lloyd. Complexity as thermodynamic depth. *Ann. Phys.*, 188:186, 1988.

39. J. E. Hopcroft and J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, Reading, 1979.

40. A. Cherubini, C. Citrini, S. Crespi-Reghizzi, and D. Mandrioli. Quasi-real-time fifo automata, breadth first grammars, and their relations. *Theo. Comp. Sci.*, 85:171 - 203, 1991.

41. E. Allevi, A. Cherubini, and S. Crespi-Reghizzi. Breadth-first context-free grammars and queue automata. *Lect. Notes in Comp. Sci.*, 324:162, 1988.

42. J. P. Crutchfield. Spatio-temporal complexity in nonlinear image processing. *IEEE Trans. Circ. Sys.*, 37:770, 1988.

43. V. I. Krinsky, V. N. Biktashev, and I. R. Efimov. Autowave principles for parallel image processing. *Physica*, 49D:247, 1991.

44. G. Mayer-Kress, editor. *Dimensions and Entropies in Chaotic Systems: Quantification of Complex Behavior*, Berlin, 1986. Springer.

45. J. P. Crutchfield and K. Young. Computation at the onset of chaos. In W. Zurek, editor, *Entropy. Complexity. and the Physics of Information*, volume VIII of *SFI Studies in the Sciences of Complexity*, pages 223 – 269, Reading, Massachusetts, 1990. Addison-Wesley.

46. J. P. Crutchfield. Reconstructing language hierarchies. In H. A. Atmanspracher and H. Scheingraber, editors, *Information Dynamics*, pages 45 – 60, New York, 1991. Plenum.

47. J. E. Hanson and J. P. Crutchfield. The attractor-basin portrait of a cellular automaton. *J. Stat. Phys.*, 66:1415 – 1462, 1992.

48. S. Wolfram. Cellular automata as models of complexity. *Nature*, 311:419, 1984.

49. S. Wolfram. *Theory and Applications of Cellular Automata*. World Scientific Publishers, Singapore, 1986.

50. T. Toffoli and N. Margolis. *Cellular Automata Machines: A New Environment for Modeling*. MIT Press, Cambridge, Massachusetts, 1987.

51. H. A. Gutowitz. Transients, cycles, and complexity in cellular automtata. *Phys. Rev. A*, 44:R7881, 1991.

52. J. P. Crutchfield and B. S. McNamara. Equations of motion from a data series. *Complex Systems*, 1:417 – 452, 1987.

53. H. Chate and P. Manneville. Spatiotemporal intermittency in coupled map lattices. *Physica D*, 32:409, 1988.

54. T. F. Meyer, F. C. Richards, and N. H. Packard. A learning algorithm for the analysis of complex spatial data. *Phys. Rev. Lett.*, 63, 1989.

55. R. Livi and S. Ruffo. Probabilistic cellular automata models for a fluid experiment. In P. Coullet and P. Huerre, editors, *New Trends in Nonlinear Dynamics and Pattern-Forming Phenomena*, volume 237 of *NATO ASI Series B*, New York, 1990. Plenum Press.

56. F. Bagnoli, S. Ciliberto, R. Livi, and S. Ruffo. Phase transitions in convection experiments. In P. Manneville, N. Boccara, G. Y. Vishniac, and R. Bidaux, editors, *Cellular Automata and Modeling Complex Physical Systems*, volume 46 of *Springer Proceedings in Physics*, Berlin, 1990. Springer-Verlag.

57. J. P. Crutchfield and K. Kaneko. Phenomenology of spatio-temporal chaos. In Hao Bai-lin, editor, *Directions in Chaos*, page 272. World Scientific Publishers, Singapore, 1987.

58. N. H. Packard. Adaptation toward the edge of chaos. In A. J. Mandell J. A. S. Kelso and M. F. Shlesinger, editors, *Dynamic Patterns in Complex Systems*, pages 293 - 301, Singapore, 1988. World Scientific.

59. C. G. Langton. Computation at the edge of chaos: Phase transitions and emergent computation. In S. Forrest, editor, *Emergent Computation*, page 12. North-Holland, Amsterdam, 1990.